



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Internalism

Citation for published version:

Ridge, M 2015, Internalism: Cui bono? in G Bjornsson, C Strandberg, R Francen Olinder, J Eriksson & F Bjorklund (eds), *Motivational Internalism*. Oxford University Press, New York, pp. 135-149.
<https://doi.org/10.1093/acprof:oso/9780199367955.001.0001>

Digital Object Identifier (DOI):

[10.1093/acprof:oso/9780199367955.001.0001](https://doi.org/10.1093/acprof:oso/9780199367955.001.0001)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Motivational Internalism

Publisher Rights Statement:

(c) Copyright Oxford University Press, 2015. All Rights Reserved.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

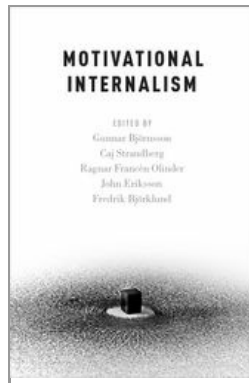
Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



University Press Scholarship Online

Oxford Scholarship Online



Motivational Internalism

Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund

Print publication date: 2015

Print ISBN-13: 9780199367955

Published to Oxford Scholarship Online: January 2015

DOI: 10.1093/acprof:oso/9780199367955.001.0001

Internalism

Cui Bono?

Michael Ridge

DOI: 10.1093/acprof:oso/9780199367955.003.0007

Abstract and Keywords

This chapter addresses the question of which meta-normative theories are best supported by the strongest form of motivational internalism that is pre-theoretically plausible. The first part of the chapter is devoted to arguing that a specific form of internalism satisfies these criteria (of strength and plausibility). The remainder of the chapter focuses primarily on how the availability of so-called “hybrid theories” in meta-normative theory makes answering the question, “which theories benefit most from the truth of internalism?” more subtle and interesting. The chapter’s conclusion is that both expressivism (hybrid and otherwise) and certain forms of hybrid cognitivism draw more dialectical dividends from the truth of the form of internalism under discussion than their rivals.

Keywords: internalism, expressivism, cognitivism, hybrid theories

So-called “judgment internalism” in metaethics and metanormative theory more generally comes in many forms.¹ What these doctrines all have in common is the assertion of some necessary connection between the relevant class of judgments (moral judgments, normative judgments more generally, or some proper subset of either) and motivation. The conventional wisdom is that the natural dialectical beneficiary of the truth of internalism is expressivism. This is because the expressivist holds that the relevant judgments just *are* motivational states. It is therefore very easy for the expressivist to accommodate even very robust forms of internalism. As it is sometimes put, the expressivist can capture internalism “for free,” whereas cognitivists must tell some special story either to accommodate internalism or explain it away.

However, so-called hybrid theories in metaethics and metanormative theory put interesting pressure on this conventional wisdom. In this chapter, I argue that it is in fact hybrid theories, whether expressivist or cognitivist, which are the most natural dialectical beneficiaries of internalism. This, in turn, complicates the debate between expressivists and cognitivist in interesting ways, or so I shall argue. I begin by laying out a version of internalism that I argue best combines plausibility and philosophical interest (section 7.1). I then explain why the availability of hybrid theories puts pressure on the idea that expressivists, in particular, are the natural beneficiaries of this form of internalism (section 7.2). I argue instead that it is hybrid theories in general that draw most benefit from internalism, so understood. I conclude by considering what form a cognitivist theory would have to take (**p.136**) to best accommodate the form of internalism on which I focus, and discuss some of the costs associated with such forms of cognitivism (section 7.3). I conclude by discussing how this way of thinking about the relevance of internalism promises to transform some of the longstanding metanormative debates.

7.1. “Goldilocks Internalism”

For my thesis about the dialectical relevance of internalism to have maximal interest, it is important to formulate internalism correctly. This is not trivial, given the wide variety of internalist doctrines by now on the market. In this section I develop a form of internalism that I call “Goldilocks Internalism,” simply to indicate my aim of striking a certain sort of balance. In particular, I want the form of internalism I discuss to be weak enough to have great plausibility and wide appeal. On the other hand, I want the form of internalism I discuss to be strong enough to be philosophically interesting in the first place.

A first question for any internalist is just what *sorts* of judgments his internalist doctrine is about. Here I focus on the *practically normative* rather than the narrowly moral. I take practically normative judgments to be ones that play a characteristic and direct role in settling the “thing to do.” Moral judgments, like the judgment that I have a moral obligation to give to charity, will be practically normative in this sense, but not all practically normative judgments will be moral. Judgments about prudential reasons, for example, are practically normative but plausibly are not moral. Perhaps judgments about distinctively spiritual and aesthetic reasons for action should also count as practically normative but non-moral.

Note that this practical/functional characterization of the practically normative need not itself beg any questions. Cognitivists can and have agreed that moral and other normative judgments play a distinctively practical role. For example, Ralph Wedgwood, takes the practical role of “ought” to be essential to its meaning, where this is cashed out in the framework of conceptual role semantics. However, Wedgwood defends a robustly Platonistic form of realism on the back of this premise about the practical role of “ought” (see Wedgwood 2007). Nor is he all that unusual in this regard; many cognitivists would be happy with some such practical/functional characterization of what I am calling “practically normative” judgments.

I prefer to couch internalism more broadly in terms of the practically normative simply because I take the arguments in favor of internalism, if sound, to be equally plausible for non-moral forms of normative judgments. Intuitively, there is something amiss in someone who judges that, all things considered, he **(p.137)** ought to exercise regularly but is not motivated to do so just as much as there is something amiss in someone who judges he morally ought to give more to charity but is not motivated to do so.

Which normative (I henceforth drop the “practically” for ease of exposition) judgments should internalism be about, though? I propose that, in the first instance anyway, the internalist should focus on verdicts. Verdicts are naturally formulated with either “ought” or “must” and contrast with judgments about what would be of some value, or what there is some reason to do. Plausibly, I can without practical irrationality judge that there is *some* reason to stay in the burning building (I can finish my delicious cake that way) even though I am not motivated in the slightest to remain there. Where a reason is, by my lights, massively outweighed, it is not at all obvious that I must be motivated even in the slightest by this reason. I therefore take the initial focus on verdicts to be more promising.

Verdicts can be expressed either with “ought” or “must” and in my view the difference matters. Plausibly, “ought” is weaker than “must” and this remains true across a wide range of contexts—epistemic as well as deontic ones, for example. “He ought to be there by now” is weaker than “He must be there by now,” where these are used epistemically. Similarly, a judgment that one morally must keep one’s promises is, in some intuitive sense, stronger than the judgment that one morally ought to call one’s mother more often. In the moral domain, this may reflect the fact that “must” is used to introduce a moral *obligation*, whereas “ought” is compatible with the action in question merely being morally good or virtuous in some way, but without necessarily thereby being morally obligatory. The action might, for example, be above and beyond the call of duty. Although it is less often invoked, I think this distinction between “ought” and “must” holds up in other deontic contexts.²

With this distinction in place, I propose that internalism is, in the first instance, much more plausible for verdicts expressed with “must” than it is for verdicts expressed for “ought.” Consider again the case of supererogation, for example. I might judge that morally I ought to give all my money to charity yet not be motivated to do so. I might judge that while I ought to do this, I am not morally obligated to do so. In that case, it is

not obvious that I am thereby irrational, though I am less than fully virtuous by my own lights.³

(p.138) More specifically, I here focus on the “all things considered must.” Judgments of this form take into account all the relevant reasons, not just (e.g.) the moral reasons. Such judgments also in some sense have as part of their content the idea that the relevant action is not just best favored by the balance of reasons, but that it is a deontic requirement. A useful contrast here might be between the judgment that there is most reason for me to eat more fruits and vegetables and the judgment that I simply must refrain from consuming hemlock.

One nice thing about focusing on the “all things considered must” is that it provides some useful traction on the so-called “amoralist” challenge to internalism. The amoralist is meant to be a character who judges that various actions are morally required, but is simply unmotivated to perform them. The intelligibility of the amoralist is meant to put pressure on various forms of internalism. The move to the “all things considered must” helps block this objection. For the way in which the intelligibility of the amoralist is made plausible, typically, is in terms of how such a character might take morality in particular to be some kind of a sham—“for suckers.” Think here of Thrasymachus from Plato’s *Republic*, or Hobbes’s *Foole*—just the sorts of characters invoked by those arguing for the intelligibility of the amoralist.

To provide a direct challenge to internalism as I have formulated it, though, one would need to defend the intelligibility of a very different character, namely the “anormativist.” The anormativist makes judgments about what she simply must do, all things considered, but is never motivated in the slightest by such judgments. Such a character is much less intuitively coherent, and certainly seems irrational in some way. In any event, the standard arguments for the intelligibility of the amoralist will not carry over neatly to the anormativist. The amoralist’s alleged intelligibility might still seem indirectly relevant, given that moral judgments plausibly are normative judgments. However, this more indirect challenge can be blocked in more ways. Perhaps the amoralist admits that moral reasons are normative but that they are always outweighed, for a start.

Much more interestingly, though, insofar as one thinks the amoralist is really intelligible, the arguments one likely would give for this view will presuppose that morality is *not* really normative. So while I agree that moral judgments are normative, there is a dialectical dilemma for the externalist here. On the one hand, she needs moral judgments to be normative if the amoralist’s intelligibility is to be as much as relevant to the intelligibility of the anormativist. On the other hand, she needs moral judgments to be non-normative insofar as without this assumption the arguments and intuitions in favor of the intelligibility of the amoralist are broken-backed.

In any event, for these sorts of reasons I think my Goldilocks Internalism is best formulated in terms of the “all things considered must.” However, not *all* verdicts couched in terms of the “all things considered must” are the best candidates **(p.139)** for an internalist treatment. For my judgments about what someone else must do need

not motivate me; more weakly, any necessary connection to motivation here will be less obvious and direct. This is most clear when it comes to my judgments about distant historical or fictional characters. *Perhaps* my judgment that Luke must not kiss Leia necessarily motivates me not to kiss Leia if in Luke's circumstances insofar as I am fully rational.⁴ This sort of indirect necessary connection to motivation is, though, much more controversial than the case of a first-person judgment couched in terms of what *I* must do, all things considered.

Should the connection between these judgments and motivation be qualified in some way? I propose that the connection be qualified in terms of the agent's *rationality*. Intuitively, it is possible to judge that I must do something and fail to be motivated to do it. That this is intuitively possible is clear because we take it to be a characteristic sort of rational failing. Someone who fails to do what they judge they must is failing by their own lights, and hence is less than fully rational. Goldilocks Internalism therefore asserts only that the relevant verdicts are such that insofar as one makes them one is either motivated accordingly or thereby irrational.⁵

What sort of motivation is necessarily connected with first-person judgments that one must act in a certain way? Plausibly, one is rationally required to *intend* to do what one judges one must do in this sense. Someone who simply had a weak desire to do what he deemed he must, but intended to do the opposite, would thereby be irrational. I therefore formulate Goldilocks Internalism in terms of intention rather than in terms of the weaker notion of having some motive, however weak, to act accordingly.

Some theorists weaken internalism by "going communal." On these views, the necessary connection is not between my judgment and my motivation, but between my judgment and my community's generally being motivated in accord with such judgments.⁶ Although this sort of internalist doctrine might also be plausible, it seems to me to be too weak to be as philosophically interesting as the more individualistic formulation I favor here. The point about the rational pressure on individuals to intend in accordance with their judgments is invisible on a communal form of internalism. Furthermore, I suspect a sort of explanatory priority in favor of the individualist formulation. In particular, my hunch is that an individualist form of internalism, plus a Davidsonian approach to interpretation, *explains* the truth of the communal formulation. However, I lack the space to defend that hunch here.

(p.140) Should the proposed necessary connection be understood as a priori or a posteriori? The latter approach would be informed by the model of "water is H₂O" where this is understood to provide an empirical but necessary truth. The analogous move would be to provide an empirical account of the nature of the relevant judgments, but to argue that this provides a necessary connection to motivation. In both cases the idea would be to provide something like an identity claim on the basis of our best empirical theory of the nature of the stuff causally regulating our discourse about it. The more traditional approach to internalism is to take the necessary connection it asserts between the relevant judgments and motivation to be an a priori and indeed an analytic one.

For present purposes, I propose to be neutral on this point, as the issue seems too controversial for a “Goldilocks” form of internalism to decide. However, to put my cards on the table, I have some sympathy for the a posteriori approach, for the following reasons. The analytic/synthetic distinction is itself rightly controversial, as is the idea of the synthetic a priori. Normative judgment is also plausibly a natural kind, and if this is so, then it is a natural candidate for the sort of semantics developed for other natural kind terms like “water,” and which are used to ground a posteriori necessities. Finally, construing internalism in this way also provides more wiggle room on the “amoralist” and even the “anormativist,” for internalism asserts a merely a posteriori necessary connection. We can now allow that such figures are indeed *conceptually* possible, but just not metaphysically possible.⁷

Finally, should internalism be understood as *de dicto* or *de re*—a useful distinction introduced by Jon Tresan (see Tresan 2006). Formally, the distinction is drawn in terms of where the necessity operator goes, roughly as follows:

- • *De Dicto*: Necessarily, if x is a normative belief then it is accompanied by suitable motivation.
- • *De Re*: Normative belief is necessarily accompanied by suitable motivation.

(p.141) The basic idea is that on a *de dicto* reading, what is being asserted is a kind of nominal essence. We just don’t count something as a moral or normative judgment without the right motivational context. However, on this approach, the very same judgment (where sameness is now understood in terms of content) could occur without the accompanying motivation. It would in that case simply not be a moral or normative judgment. Whereas on a *de re* reading, it is the nature of the judgment itself—its real essence—which essentially precludes it from occurring without the associated motivation. It is useful to compare the necessity of planets orbiting stars, which is a *de dicto* necessity. The very same celestial object could exist without orbiting a star, but it would then not count as a planet. Tresan argues that we should understand internalism in this *de dicto* way, rather than *de re*.

I, however, prefer the more traditional (if often implicit) *de re* reading. This is, admittedly, a stronger form of internalism, but I think it provides a correspondingly more interesting doctrine. Insofar as the distinction matters, moreover, the *de dicto* version looks dubious.

First, plausibly, normative beliefs are necessarily normative beliefs. The contrast with wishful thinking or inspirational beliefs (two of Tresan’s models) is striking. Plausibly, judgments are individuated, at least in part, by what goes into their “that” clauses. Notice how weird the following claim is: “Yesterday he believed that eating meat is wrong. Today, he has that very same belief, but he no longer believes eating meat is wrong.”

Second, the *de dicto* reading does not fit very well with the form of internalism I am after. For I want to capture the idea that failing to be motivated is possible, but irrational. To

shoehorn this form of internalism into a *de dicto* model would require that we make the relevant category not simply “such-and-such normative belief” but rather “such and such normative belief as held by a fully rational agent.” It is simply implausible that such a gerrymandered concept as this has a merely nominal essence in the way Tresan’s *de dicto* internalism as adapted here would presuppose.

Here, then, at last, is my proposed formulation of Goldilocks Internalism:

Necessarily, for any agent A, If A is practically rational and makes a first personal judgment [that is, with the indexical ‘I’] that A must Φ in C, all things considered, then A’s judgment is necessarily accompanied by an intention to Φ in C.⁸

(p.142) I now turn to the question of what theories draw the most natural dialectical benefits from the truth of Goldilocks Internalism.

7.2. Cui Bono?

The traditional view is that expressivism is the natural dialectical beneficiary of internalism. This is not surprising, since the practicality of normative judgment is often invoked as one of the main arguments for expressivism. Moreover, expressivists do get internalism “for free” given their theory. For if normative judgments just *are* motivating states, then trivially normative judgments will be necessarily associated with motivating states. By contrast, cognitivists who want to accommodate internalism will need to tell some fancy story about the content or character of normative judgments in order to explain how a representational state like a belief might nonetheless be necessarily connected with motivation. Moreover, those who find a broadly Humean philosophy of mind, according to which (a) beliefs and desires are “distinct existences,” and (b) belief as such is motivationally inert, will find it unlikely, to say the least, that a cognitivist could successfully accommodate internalism.

However, this traditional view is mistaken. For there are *hybrid* forms of cognitivism which (a) are equally compatible with a Humean philosophy of mind, yet (b) can accommodate internalism on the cheap. Hybrid theories come in both cognitivist and expressivist forms, but in my view hybrid forms of cognitivism *and* hybrid forms of expressivism do better in terms of accommodating internalism than their non-hybrid cousins. Let me begin with a discussion of how certain hybrid forms of cognitivism are at least as well placed as non-hybrid forms of expressivism to accommodate internalism.

Hybrid theories in metanormative theory in the sense invoked here hold at least one of the following theses:

- (1) Normative claims express *both* belief-like *and* desire-like mental states.
- (2) Normative judgments are either constituted by or individuated in terms of both belief-like and desire-like components.

Hybrid theories can come in both cognitivist and expressivist forms. I define cognitivism and expressivism as follows:

- • Cognitivism: For any normative claim *N*, *N* expresses a [representational] belief such that *N* is thereby guaranteed to be true if and only if the belief is true, where the content of the belief expressed is the same as the content of the claim expressing it.
- **(p.143)** • Expressivism:
 - (1) For any normative claim *N*, *N* does not express a [representational] belief such that *N* is thereby guaranteed to be true if and only if the belief is true. (non-cognitivism)
 - (2) For any normative claim *N*, *N* expresses a state which is either desire-like or necessarily constituted in part by a desire-like state.

I have added the clause “representational” in the definition of expressivism because expressivists often allow that there is a deflationist sense in which normative judgments are beliefs. However, in the sense invoked by Humean theories in the philosophy of mind, normative judgments are not, according to the expressivist, well understood as beliefs (qua representational states) with distinctively normative contents.

A *hybrid* form of cognitivism then adds to the definition of cognitivism proposed above. either the thesis that normative claims also express desire-like states in some sense *or* the thesis that normative judgments are partly constituted by or individuated in terms of desire-like states.

Hybrid forms of cognitivism can therefore be further divided in theoretically important ways. Most notably, some hybrid cognitivists allow a role for desire only at the level of *pragmatics*. On these views, normative sentences are semantically just like descriptive sentences (apart from the obvious differences in content). However, speech-acts in which such sentences are used to make assertions also pragmatically convey the speaker’s desire-like state(s) in some way. A standard move here is to appeal to the Gricean idea of “implicature” and to argue that normative claims implicate a speaker’s relevant desire-like states. Such implicatures views can be further divided in terms of whether they take the relevant implicatures to be conversational or conventional.

However, such implicature views work only at the level of public language. Yet the internalism is a doctrine about normative thought. Because the relevant pro-attitudes are not on these views any part of the judgments themselves, nor rationally demanded by them (at least, not in virtue of the pragmatic story), they do not accommodate internalism by such implicatures. Defenders of implicative views have things to say at this point, and some of them aim to debunk or explain away internalist intuitions rather than to accommodate internalism. However, my aim here is not to assess the overall plausibility of such views, but simply to see how far they can go in terms of explaining Goldilocks Internalism.

More robust hybrid forms of cognitivism *do* operate at the level of normative thought rather than just at the level of public language. On these accounts, normative judgments are *individuated* in terms of the presence of the relevant pro-attitudes. Someone simply does not count as making the relevant judgment unless she also is motivated in some way

specified by the theory.

(p.144) Such “judgment-individuating” views can be further divided according to whether they take the relevant desire-like state to literally partly constitute the normative judgment. If they do, then they hold a *constitutive* form of hybrid cognitivism. If they don’t, then they hold a *con-constitutive* but still *judgment-individuating* form of hybrid cognitivism. The latter view can draw inspiration from an approach like Tresan’s, according to which the very same judgment can occur without the motivation in question but will then simply not count as normative. On this view, the motivational state in question is not naturally thought of as literally being part of the judgment, so much as a necessary condition on the judgment’s counting as normative. Because this approach presupposes that internalism is analytic, though, I prefer constitutive forms of judgment-individuating hybrid cognitivism. Recall that I suggested in section 7.1 that it is more plausible to suppose that internalism is a posteriori than analytic. However, not much really hangs on this here. My primary point is that cognitivists need to go beyond merely implicative views to accommodate Goldilocks Internalism. Whether this is best done by adopting a constitutive or non-constitutive form of judgment-individuating hybrid cognitivism is not so important, really.

These views really do look well poised to accommodate Goldilocks Internalism, though. They can hold that for me to count as making a normative judgment of the form “I must Φ ” that I must intend to perform actions of a certain kind K and at the same time believe that my Φ -ing would be K. K, then, will provide the semantic content for the “all things considered must,” but for someone to count as making a judgment of the form “I must Φ in C,” it is not enough that they believe that their Φ -ing in C would be K. They must also intend to perform actions that are K.

So long as we can help ourselves to the very plausible assumption that it is practically irrational to intend to perform actions of a given kind, believe that this action would be of that kind, yet fail to intend to perform this action, a view with this structure can accommodate Goldilocks Internalism. Moreover, it accommodates Goldilocks Internalism “on the cheap.” There is no pressure on this sort of view to abandon a Humean philosophy of mind, or to tell some special story about normative judgment as a distinctive and sui generis yet unified mental state which is at once representational and motivational. Rather, a role for motivation is built into the theory of the nature of the relevant judgments at the ground level, and in a way which does not presuppose the falsity of Humean views in the philosophy of mind.

So cognitivists who “go hybrid” can steal the expressivist’s thunder when it comes to the motivating power of normative judgment. However, there is something very even-handed here, in that (as I have argued elsewhere; see Ridge 2006) expressivists who “go hybrid” can steal the cognitivist’s thunder **(p.145)** on a number of fronts, perhaps most notably in terms of more easily meeting the challenge poised by the notorious Frege-Geach problem. Nor should this be surprising. Many of the puzzles of metanormative theory arise out of the apparently Janus-faced nature of normative judgment. By including both a belief-like and a desire-like component in their theories at

the ground level, both expressivists and cognitivists stand a better chance of adequately accommodating this Janus-faced nature. The expressivist does better at accommodating belief-like features by going hybrid, while the cognitivist can do better at accommodating the desire-like features of normative judgment by going hybrid.

In fact, expressivists can actually do better when it comes to accommodating internalism by “going hybrid.” In particular, they can avoid being committed to a form of internalism that is implausibly strong, and which precludes their offering a compelling account of akrasia by going hybrid. Let me explain.

On Allan Gibbard’s preferred form of expressivism, judging I ought to (or must) Φ in C just is planning to Φ in C. Trivially, then, there will be no logical gap between, for example, judging that I must Φ and my planning to Φ . Given that planning is just a form of intending, though, this means there is no logical gap between judging I must Φ in C and intending to do so. To return to my Goldilocks metaphor, Gibbard’s internalism isn’t “just right”—it is too strong. Intuitively, it is possible to engage in what is sometimes called “clear eyed akrasia,” in which one knowingly acts against one’s normative judgment while still holding that judgment.

In fairness, Gibbard owns up to this implication of his view and tries to argue that it is not so implausible as it seems. There is a tradition, going back to Socrates in the early dialogues, of defending such a strong connection between judgment and motivation. I lack the space to discuss the various moves and countermoves here. Instead, I just want to call attention to the fact that this feature of Gibbard’s view is no accident. For insofar as one insists that normative judgment just is a motivating state of the right sort, one will fail to be able to make the right kind of sense of this specific form of akrasia—akrasia in which one simply does not form the corresponding intention, as opposed to akrasia in which one does intend appropriately but is overwhelmed by a stronger desire, caves into temptation, and does not follow through on one’s intention. Gibbard’s expressivism can accommodate the latter form of akrasia but not the former.

A hybrid form of expressivism can do better. On a simple version of my own view, which I call “Ecumenical Expressivism,” a normative judgment might be a *general* plan to perform actions of a given type and a belief that this action is of that type. However, it is well known that one’s general intentions can fail to issue in immediate/proximate intentions. Clear-eyed akrasia (**p.146**) is possible on this view, precisely in virtue of the gap between general and specific motivation. Moreover, this can also explain why such akrasia is irrational. For such akrasia will, on this account, consist in a general intention to do whatever actions have a certain property, a belief that this action has precisely that property, yet a failure to intend to perform this action. So Goldilocks internalism favors hybrid or “Ecumenical” Expressivism if it favors any kind of expressivism at all.

So both cognitivists and expressivists can give a better explanation of Goldilocks Internalism “on the cheap” by going hybrid. In the last and final section, I explore what sorts of further constraints might be imposed on a cognitivist who wants to “go hybrid” in this way in order to explain Goldilocks Internalism.

7.3. What Form of Hybrid Cognitivism?

I have already argued that hybrid cognitivists should favor some form of “judgment individuating” hybrid cognitivism to accommodate internalism. In this section I consider what further constraints might plausibly be imposed on such a strategy. I do not explore the analogous question for hybrid expressivism simply because I have developed my own version of that view in detail elsewhere, and the details of that theory indicate what I take to be some of the most important constraints on any such theory.

One natural question at this stage is whether there are any further constraints imposed on the *content* of the “all things considered must” in virtue of the sort of hybrid cognitivist theory canvassed in the previous section. While no such further constraints logically follow from the hybrid approach, there are some minimum constraints which are very plausible. Perhaps most minimally, the content should be such that actions which satisfy that content are within the agent’s power. Otherwise we will have “must” without “can”, which seems dubious.

More interestingly, it may be that *non-naturalist* accounts of the normative do not fit well with this approach. Indeed, those who have defended such hybrid theories in the literature have all favored some form of naturalism, and I do not think this is a coincidence. Let me explain.

Michael Smith has argued, plausibly to my view, that externalist accounts of normative motivation make the virtuous and strong-willed person out as fetishizing rightness. This is because the externalist must attribute a *de dicto* desire to do the right thing, whatever that turns out to be. Otherwise the externalist cannot explain why the virtuous and strong-willed person’s motivations systematically change over time in line with changes in their judgments. (p.147) Whatever one thinks of this argument (and it is controversial), anyone sympathetic with ambitious forms of internalism should find it congenial. It is, after all, a powerful argument in favor of internalism. So the dialectic makes this not unreasonable to assume. Moreover, it is not all that implausible to suppose that if the argument works at all, it works in the normative realm tout court as well as for narrowly moral judgments. Though given that on my view moral judgments are at least a species of normative judgment, this may not matter all that much—the fetishism argument can still be deployed so long as the hybrid cognitivist doesn’t take a different view in the specifically moral case.⁹

This has interesting consequences for the hybrid cognitivist who wants to accommodate Goldilocks Internalism, though. It seems to suggest that *anti-reductive* forms of cognitivism, whether naturalistic or non-naturalistic, will run into problems if they try to “go ecumenical” to *explain* internalism. The problem for anti-reductionism: the only content that seems to ensure the right diachronic motivational profile would be something like “that I do what I must.” Given an anti-reductive theory, there is no specific *descriptive* content for normative claims, after all. That, though, is just the sort of motive Smith’s critique targets.

Of course, there will, trivially, be descriptively specifiable truth-conditions for normative

claims, even given an anti-reductive view. This is trivially true in virtue of the supervenience of the normative on the descriptive. However, these truth-conditions might be massively complex and disjunctive for all that has been said so far. Indeed, non-naturalists often favor a kind of radical pluralism or even “particularism” on the strength of their non-naturalism (here think of Ross, historically, and Jonathan Dancy, more recently).

Even if, as Moore thought, the descriptive truth conditions are much simpler than this, though, it would be dialectically odd to import these truth conditions into a hybrid theory of what one must desire to count as making the relevant normative judgment, for anti-reductive views like Moore’s are driven by the so-called “Open Question Argument.” That argument takes it that one’s normative stance is never fixed, simply qua conceptual and semantic competence, by any descriptive take on the world. To insist that there is some specific descriptive content, such as maximizing happiness, that one must desire in order to count as making a normative judgment would seriously contradict the motivations underlying such anti-reductive views in the first place. Such a view would hardly be attractive.

So if one wants to accommodate, rather than reject, Goldilocks Internalism “on the cheap” as a cognitivist, then one ought to be a reductionist cognitivist. That is in itself an interesting result, and perhaps casts interesting light (p.148) on the fact that hybrid cognitivists in the literature have without exception favored reductionism. A reductionist cognitivist can, of course, without any awkwardness of the sort just discussed for the anti-reductionist, simply take the descriptive content their theory assigns to the normative “must” and then hold that one must have a desire with that content in order to count as making a normative judgment in the first place.

7.4. Conclusion

I have here argued that both cognitivists and expressivists can better accommodate what I take to be the most philosophically interesting form of internalism (Goldilocks Internalism) on the cheap by “going hybrid.” I have further argued that cognitivists who take this strategy should go for a *constitutive*, rather than an *implicative* form of hybrid theory, and that they need to be reductionists about the semantic content of normative predicates.

Since my own view is a form of hybrid or “Ecumenical” Expressivism, one might wonder how I think all of this bears on the larger debate between myself and my cognitivist hybrid cousins. One point worth noting here is that reductionism plus *Constitutive* Ecumenical Cognitivism makes the very strong prediction that anyone who has ever made a normative judgment thereby has a desire for some specific naturalistic property or state of affairs—maximizing happiness, or whatever. That simply seems implausible on its face.¹⁰

In my view, reductionism also leads them into trouble in terms of Moore’s “Open Question Argument” and what I think of as its successors—namely R. M. Hare’s parable of the missionary and the cannibal and Horgan and Timmons on normative/moral twin

earth. These points are not unrelated, since moral twin earth examples in effect get their force from the intuition that normative disagreement is a kind of fundamental practical disagreement rather than some sort of factual disagreement. However, explaining exactly how all this bears on hybrid forms of cognitivism is a story for another day.¹¹

References

Bibliography references:

Bedke, M. 2009. "Moral Judgment Purposivism: Saving Internalism from Amoralism." *Philosophical Studies* 144: 189–209.

Björnsson, G. 2002. "How Emotivism Survives Immoralists, Irrationality and Depression." *Southern Journal of Philosophy* 60: 327–344.

Gibbard, A. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press

Ridge, M. 2006. "Ecumenical Expressivism: Finessing Frege," *Ethics* 116: 302–336.

Ridge, M. 2014. *Impassioned Belief*. Oxford: Oxford University Press.

Schroeder, M. 2009. "Hybrid Expressivism: Virtues and Vices," *Ethics* 119: 257–309.

Smith, M. 1994. *The Moral Problem*. Oxford: Basil Blackwell.

Tresan, J. 2006. "De Dicto Internalist Cognitivism." *Nous* 40: 143–165.

Wedgwood, R. 2007. *The Nature of Normativity*. Oxford: Oxford University Press.

Notes:

(1) Many thanks to Caj Strandberg for detailed comments on a previous draft of this material, and to the participants of the conference in Gothenburg who also provided many helpful comments.

(2) In work in progress, I distinguish "ought" from "must" quite generally (and not just in deontic contexts) in terms of a distinction between standards which recommend and standards which require, where the context determines the kind of standards being invoked. See Ridge (2014).

(3) The discussion of whether supererogation might bring other motivational or affective states in its wake, as a matter of necessity, has not been much discussed. One interesting candidate here might be admiration, or a motive to praise. In his forthcoming PhD dissertation, Alfred Archer discusses this issue in some detail.

(4) Compare Gibbard (2003).

(5) This is not an uncommon view; compare, e.g., Smith (1994).

(6) See, e.g., Tresan (2006).

(7) Cf. Bedke (2009) and Björnsson (2002) for some discussion of how this might go. My own view is actually more complex on this front. I think that the internalist thesis I put forward in the text (Goldilocks Internalism) actually can be defended on a priori grounds. However, although I lack the space to discuss this here, I think that Goldilocks Internalism can then be used as a premise in an argument for another form of internalism that can itself be established partly on empirical grounds. This other form of internalism drops the “or you are irrational clause” but weakens the connection to motivation to a mere capacity to be motivated accordingly. Some of the additional empirical facts needed to establish this stronger form of internalism are empirical facts about the ubiquity of deep normative disagreement. One useful framework here is that both a priori constraints and empirical facts can help us determine the best theory of the states that play the role of normative judgments for us, so that both a priori and empirical facts can be relevant to determine what realizes that role. I discuss these nuances more in Ridge (2014).

(8) Note that this formulation includes both a wide scope and a narrow scope necessity operator. This reflects the fact that I think *de re* internalism should itself be understood as a necessary truth, if it is true at all. This marks a contrast with the formulation drawn from Tresan above, but one which preserves the *de re* reading. The thesis would be *de dicto* only if there were no narrow scope necessity operator since that is what ensures that the relevant judgment is necessarily accompanied by motivation (given practical rationality, on my formulation).

(9) Thanks to Caj Strandberg for drawing me out on this point.

(10) Cf. Mark Schroeder’s exchange with Daniel Boisvert, where Schroeder argues that insofar as this is true it makes the fancy hybrid machinery somewhat otiose and unmotivated (Schroeder 2009).

(11) I try my hand at it in chapter 3 of Ridge (2014).



Access brought to you by: University of Edinburgh